DOCUMENT RESUME

ED 279 973 CG 019 782

AUTHOR Futterman, Robert; Jones, Judith
TITLE Methodological Issues Related to the Evaluation of Comprehensive School-Based Clinics.
SPONS AGENCY Carnegie Corp. of New York, N.Y.; Community School District 6, New York, N.Y.; Hewlett Foundation, Inc., Garden City, N.Y.
PUB DATE 30 Sep 86
NOTE 13p.; Paper presented at the Annual Meeting of the American Public Health Association (114th, Las Vegas, NV, September 28-October 2, 1986).
PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Clinics; Decision Making; Elementary Secondary Education; Evaluation Criteria; *Evaluation Methods; *Program Evaluation; *Research Problems; *School Activities; *Validity

ABSTRACT

As the number and scope of school-based clinics grow, interest in their evaluation increases. Evaluation of a comprehensive school-based clinic program in New York City revealed methodological issues in evaluation of these programs. The first issue involves program goals and objectives. Even before program implementation, the evaluator should review goals and objectives and determine whether program objectives are logically related to goals. Evaluators of school-based clinics should adopt the decision-making evaluation model which is oriented to the informational needs of administrators and policymakers. The most persuasive evidence about the success of school-based clinic programs will come from evaluations which use experimental or quasi-experimental research designs. The use of comparison groups is one method of controlling for the impact of unforeseen historical events which may be mistaken for program effects. Process measures are useful for linking program operations to program impacts. During data analysis, it may be helpful to either statistically adjust for the effect of subject selection or analyze equivalent subsamples of students. An evaluation design must also control for attrition in estimating program effectiveness. Evaluating with validity involves the choice of appropriate process and outcome measures. Intervening variables, the processes by which the program produces its effects, should also be measured. Microcomputers can aid in data collection, analysis, and information integration. In order to provide a useful and valid evaluation, theoretical and practical issues must be addressed. (NB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED279713

Methodological Issues Related to the Evaluation of Comprehensive School-Based Clinics

Robert Futterman, Ph.D.*
Judith Jones. M.S.*

*Faculty of Medicine of Columbia University,
Center for Population and Family Health

Address Reprint Requests to:

Robert Futterman
Associate Research Scientist, CPFH
60 Haven Avenue, B3
New York, NY 10032

CG019782

Paper to be presented at the Annual Meeting of the American Public Health Association, September 30, 1986, Las Vegas, Nevada

Acknowledgments: The research reported on in this paper was made possible by grants from the Carnegie Corporation and the Hewlett Foundation, with the cooperation of Community School District 6 in New York City, Loyda Alfalla, District Superintendent.

Please do not cite or quote without permission from the authors.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Robert Futterman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

METHODOLOGICAL ISSUES RELATED TO THE EVALUATION OF COMPREHENSIVE SCHOOL-BASED CLINICS

Robert Futterman, Ph.D., Associate Research Scientist
Judith Jones, M.S., Associate Clinical Professor
Center for Population and Family Health, Columbia University

As the number and scope of school-based clinics across the nation grows, interest in their evaluation correspondingly increases. This paper examines some methodological issues related to the evaluation of a comprehensive school-based clinic program in New York City. I will describe some of the problems we have encountered and some of the solutions we have devised.

The first evaluation issue concerns the development of goals and objectives. The evaluator should review the program's goals and objectives and determine whether program objectives are logically related to the program's goals. Thus, if the program goal is to reduce unintended teen pregnancy, the program's objectives must be directed toward that goal and not to toward more subordinate goals such as reducing substance abuse or other dysfunctional behaviors. A statement of program objectives should be written which specifies proposed program activities and quantitative criteria for measuring whether the objective has been achieved. Ideally, the program design should be based on a conceptual model which describes the links between the problems that the school-based clinic intends to address, the proposed program inputs and program processes, and the desired outcomes of the program. By critiquing the statement of goals and objectives

and program model, the evaluator has the opportunity to improve the program before implementation begins.

The first issue in designing an evaluation is the choice of an evaluation approach or model. One implication of choosing an evaluation model is that the kinds of evidence that are most important to your audiences must be collected and reported. The evaluation's potential audiences and the types of evaluation information they may require or desire must be considered. It is critical to prepare a comprehensive evaluation design early in the program planning period which specifies the evaluation questions that will be addressed and the types of information that will be collected. The evaluation design must be discussed with program managers, staff, and other evaluators if warranted, and revised if necessary.

As Figure 1 shows, there are numerous approaches to evaluation. Evaluation models vary according to major audiences or reference groups, assumptions, methodology, outcomes, and typical questions. We suggest that evaluators of school-based clinics adopt the decision-making evaluation model. As the name implies, the decision-making approach to evaluation is oriented to the presumed informational needs of policy-makers and administrators. For managers and administrators of school-based clinics, the primary evaluation question is likely to be, "Did the program work?". Without convincing evidence that the program is effective, other questions (for instance, questions concerning cost efficiency or replicability) are of secondary importance.

These other concerns should be investigated, but not at the risk of neglecting to acquire information about program effectiveness.

We believe that the most persuasive evidence about the success of school-based clinic programs will come from evaluations which employ experimental or quasi-experimental research designs, because they control for numerous threats to the validity of the evaluation. By validity, I am referring to the confidence in the evaluation's findings about program effectiveness. Some of the major threats to validity include history, attrition, and selection. Our evaluation design attempts to control for these and other threats to validity within logistical and financial constraints. Although the most thorough evaluation design would incorporate the simultaneous measurement of treatment and comparison groups at every measurement point, our evaluation design specifies that comparison group students will be measured only twice -- at the baseline and again at the end of the evaluation -- rather than annually. This decision demonstrates how the scope of the evaluation effort must be in proportion to the program's financial and personnel resources.

History becomes a threat to validity when unanticipated events which occur during the evaluation are mistaken for program effects. The schools may offer a new sex education course, a health clinic may be established in the community, or a clinic school may be closed down. The use of comparison groups is one method of controlling for the impact of unforeseen historical events. Without a comparison group, it is impossible to determine whether outcomes are due to the program or to these

unanticipated events. Treatment and comparison groups should be drawn from the same locale or administrative jurisdiction whenever possible. If this is done, it is more likely that local historical events will affect both the treatment and control groups. The evaluation we have designed includes the annual measurement of all students in schools with clinics as well as all students in a similar school without a clinic. Another method we are using for examining the effects of historical events is through documentation of the process of program implementation, using project memoranda; clinic records including sampling of medical charts; logs and appointment books; and even diaries kept for this purpose by program staff. This qualitative information greatly enriches understanding of the program and its effects.

In addition, process measures allow program evaluators to link program operations to program impacts. They also provide funders, administrators, replicators and program analysts with a description of the program. Qualitative process measures such as interviews with clinic or school staff should be systematically collected, in addition to quantitative process measures such as number and type of clinic encounters. Both the extent and quality of services should be documented. Although process measures are particularly valuable in formative evaluation, they should also be employed during summative evaluation because they provide essential contextual information for interpreting quantitative results.

Simply including comparison groups in the evaluation design does not insure validity. The greater the pretest differences between the clinic and comparison students, the less confidence one has that posttest differences between those groups are actually due to the effects of the program. In other words, biases in the selection of these groups can seriously threaten internal validity. In an experimental setting, students would be randomly assigned to the treatment and control groups to eliminate selection bias. In most school-based clinic programs, randomized access to clinic services is not possible for both practical and ethical reasons. However, any process that can be randomized -- for instance, the assignment of clinics to schools, grades, or groups -- ought to be randomized in order to minimize selection bias.

Failing this, the extent of selection bias should be investigated by collecting thorough information on variables which characterize the treatment and comparison groups. During data analysis, it may be possible either to statistically adjust for the effect of selection (using analysis of covariance) or to analyze equivalent subsamples of students (for instance, black or Hispanic students) rather than the entire sample. As indicated earlier, whenever control over the selection process is possible, comparison groups should be formed of students, classes, or schools which have the greatest similarity to each other. Relevant variables would include ethnic composition, academic achievement levels, or economic status. This information should be available from the school system.

Another issue that impacts on the strength of the evaluation is program attrition. Attrition reduces statistical power and confidence in statistical conclusions. For instance, if pregnant teens drop out of school at a higher rate than nonpregnant teens, then the program may mistakenly seem to reduce teen pregnancy rates. An evaluation design which does not attempt to control for attrition may be unable to accurately estimate program effectiveness. In school districts such as the one in which our program is located, student mobility rates are very high -- over 35 percent of the student register changes during the course of the school year. To minimize the effect of attrition on the evaluation's validity, it may be necessary to identify individual student responses while maintaining their confidentiality. This would permit longitudinal data analyses limited to student who have remained in the program. Student identification is also necessary to conduct follow-up studies of treatment and comparison students. To maximize statistical conclusion validity, the data collection process must be thoroughly controlled. Response rates should be acceptably high (80% or higher) and program attrition should be monitored and documented.

Attrition is an even more severe threat to validity when there is differential attrition in the treatment and comparison groups. Nonetheless, the determinants of attrition can be examined using various statistical techniques. In one recent study a probit model was used to estimate if group membership (treatment or comparison) was a determinant of attrition. It might be best to allocate additional financial or staff resources

to collect data from those students most likely to drop out of the program. They may already be earmarked by school officials, as in New York City, where about 200 youths in each of the junior high schools in which our program operates are part of the Attendance Improvement Dropout Prevention (AIDP) program.

Evaluating with validity also involves the choice of appropriate process and outcome measures. The success of the clinic will probably be judged on the basis of a relatively few measures including: clinic use, delay of initiation of first intercourse, client satisfaction and compliance with medical regimens. Conceptualizing and operationalizing these measures deserves the greatest amount of care. Measures of established validity and reliability (such as items from Kanter and Zelnick's surveys) should be employed when possible -- these are useful for comparing the impact of one program to another. Comparability of measures across projects permits a common metric for judging project effectiveness.

When using previously-developed measures, bear in mind the cognitive capabilities of your clients; instruments designed for adults or students with average achievement levels will not be appropriate for children or for students with low achievement levels. The response format as well as the item stems should be reviewed for comprehensibility. Our experience with Hispanic students has made us aware of the need to translate all materials and review the translations with local residents who are native Spanish speakers. For all these reasons, it is essential to pretest measures before committing oneself to their use.

Outcome measures must have a logical relationship to project goals and objectives. If, for example, the program is designed to reduce the incidence of teens' sexual behaviors, then outcome measures should not be restricted to students' knowledge of reproductive physiology. Rather, they should include measures of delay in initiation of intercourse, contraceptive utilization, etc.

The processes by which the program produces its effects should also be measured. These intervening variables may be psychological variables such as tolerance to deviance, sociological variables such as socioeconomic status or process variables such as the number of clinic encounters per student. The type and extent of program exposure is another conceptually important intervening variable which should be measured for each student.

There are numerous sources of evaluation information, as well as methods to collect this information. For instance, we intend to collect information from all students using a self-administered questionnaire because the logistical difficulties in selecting a sample of students to receive the questionnaire and the cost of employing professional interviewers deterred us from designing a sample survey. Student health information will be collected from the clinic records, academic information will come from the school records, service providers will be interviewed, and the clinics will be systematically observed. Diversifying informational sources and methods is likely to increase validity.

Microcomputers can aid in data collection, analysis and information integration. We have installed microcomputers (PC-XTs) in our health clinics for these purposes. We will use a hierarchical data-base management software program (SIR-PC) to design forms for computerized data entry. Student health information can be stored and analyzed at the clinic site using this software. Additionally, this data can be easily converted into a SPSS-X systems file for sophisticated data analyses using either mainframe or microcomputers.

In sum, methodological issues are of great importance in evaluating school-based clinics. Both theoretical and practical issues are relevant if the evaluation is to prove both valid and useful to researchers, practitioners, administrators, and funders.

FIGURE 1 A Taxonomy of Major Evaluation Approaches

Model	Major Audiences or Reference Groups	Assumes Consensus On	Methodology	Outcome	Typical Questions
Systems analysis	Economists, managers	Goals, known cause and effect, quantified variables	PPBS, linear programming, planned variation, cost-benefit analysis	Efficency	Are the expected effects achieved? Can the effects be achieved more economically? What are the most efficient programs?
Behavioral objectives	Managers, psychologists	Prespecified objectives, quantified outcome variables	Behavioral objectives, achievement tests	Productivity, accountability	Is the program achieving the objectives? Is the program producing?
Decision-making	Decision-makers, especially administrators	General goals, criteria	Surveys questionnaires, interviews, natural variation	Effectiveness, quality control	Is the program effective? What parts are effective?
Goal-free	Consumers	Consequences, criteria	Bias control, logical analysis, modus operandi	Consumer choice, social utility	What are all the effects?
Professional review	Professionals, public	Criteria, panel, procedures	Review by panel, self study	Professional acceptance	How would professionals rate this program?

Adapted from Ernest House, Evaluating with Validity. Sage, 1980

BEST COPY AVAILABLE